

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MP-etch-pp protein - protein database search, using Smith-Waterman algorithm

Thu Jan 27 17:33:59 2000; Maspar time 15.52 Seconds

496.003 Million cell updates/sec

Tabular output not generated.

Title: >US-09-397-558-1
Description: (1-141) from US09397558.pep
Perfect Score: 1097
Sequence: 1 MVHVAVSLCLPMBRSERYLF.....ALVPSIVLIDLLQCRYPD 141

Scoring table: PAM 150
Gap 11

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb19
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 44.978; Variance 132.607; scale 0.339

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	125	11.4	539	5	F40F9.5 PROTEIN.	9.36e-02
2	121	11.0	552	2	PUTATIVE PHOSPHATE PER	2.32e-01
3	120	10.9	380	8	CYTOCHROME B.	2.91e-01
4	116	10.6	488	8	NADH DEHYDROGENASE SUB	7.08e-01
5	116	10.6	1055	3	P-TYPE ATPASE 1.	7.08e-01
6	115	10.5	472	2	HOMOLOGUE OF MULTIDRUG	8.83e-01
7	112	10.2	256	14	HYPOTHETICAL 29.9 KD P	1.70e+00
8	112	10.2	386	2	HYPOTHETICAL 43.8 KD P	1.70e+00
9	110	10.0	205	8	YCF15 PROTEIN.	2.61e+00
10	110	10.0	319	5	F09P3.1 PROTEIN.	2.61e+00
11	109	9.9	381	8	CYTOCHROME B.	3.24e+00
12	109	9.9	695	8	NADH DEHYDROGENASE (FR	3.24e+00
13	107	9.8	125	2	HYPOTHETICAL 13.1 KD P	4.95e+00
14	108	9.8	192	8	CYTOCHROME B (FRAGMENT	4.01e+00
15	108	9.8	300	8	CYTOCHROME C (FRAGMENT	4.01e+00
16	108	9.8	311	8	CYTOCHROME C (FRAGMENT	4.01e+00
17	107	9.8	311	8	CYTOCHROME C (FRAGMENT	4.95e+00
18	108	9.8	379	8	CYTOCHROME B.	4.01e+00
19	108	9.8	381	8	CYTOCHROME B.	4.01e+00
20	108	9.8	381	8	CYTOCHROME B LIGHT STR	4.01e+00

21	107	9.8	381	8	P92717	CYTOCHROME B.	4.95e+00
22	107	9.8	381	8	C35545	CYTOCHROME B.	4.95e+00
23	107	9.8	414	2	007867	PUTATIVE POLYSACCHARID	4.01e+00
24	108	9.8	475	2	066589	NITRATE TRANSPORTER	4.01e+00
25	108	9.8	675	8	032213	NADH DEHYDROGENASE SUB	4.01e+00
26	107	9.8	685	8	019822	NADH DEHYDROGENASE SUB	4.95e+00
27	107	9.8	696	8	032058	NADH DEHYDROGENASE (FR	4.01e+00
28	108	9.8	702	8	019823	NADH DEHYDROGENASE (FR	4.01e+00
29	106	9.7	192	8	035642	CYTOCHROME B (FRAGMENT	6.11e+00
30	106	9.7	368	1	058546	368AA LONG HYPOTHETICA	6.11e+00
31	106	9.7	379	8	P92657	CYTOCHROME B.	6.11e+00
32	105	9.7	380	8	047545	APOCYTOCHROME B.	6.11e+00
33	105	9.7	380	8	019831	NADH DEHYDROGENASE (FR	6.11e+00
34	106	9.7	684	8	019825	NADH DEHYDROGENASE (FR	6.11e+00
35	106	9.7	685	8	019854	NADH DEHYDROGENASE-LIK	6.11e+00
36	106	9.7	699	8	019856	NADH DEHYDROGENASE-LIK	6.11e+00
37	106	9.7	699	8	019853	NADH DEHYDROGENASE-LIK	6.11e+00
38	106	9.7	699	8	019855	NADH DEHYDROGENASE-LIK	6.11e+00
39	106	9.7	702	8	019830	NADH DEHYDROGENASE (FR	6.11e+00
40	106	9.7	710	8	019824	NADH DEHYDROGENASE (FR	6.11e+00
41	105	9.6	197	1	028693	HYPOTHETICAL 21.8 KD P	7.54e+00
42	105	9.6	313	5	045426	F31E9.2 PROTEIN.	7.54e+00
43	105	9.6	348	8	095935	CYTOCHROME B (FRAGMENT	7.54e+00
44	105	9.6	382	8	035047	CYTOCHROME B LIGHT STR	7.54e+00
45	105	9.6	382	8	035518	CYTOCHROME B LIGHT STR	7.54e+00

ALIGNMENTS

RESULT	ID	Q20366	PRELIMINARY;	PRT;	539 AA.
AC	Q20366	(1-NOV-1996 (TREMBL:REL. 01, CREATED)			
DT	11-NOV-1996 (TREMBL:REL. 01, LAST SEQUENCE UPDATE)				
FE	01-JAN-1999 (TREMBL:REL. 09, LAST ANNOTATION UPDATE)				
DE	F40F9.5 PROTEIN.				
CS	F40F9.5.				
CC	CAEORHABDITIS ELEGANS.				
GG	ENKAROTA: METAZOA: NEMATODA: SECERNENTEA: RHABDITIDA: RHABDITIDA:				
CC	RHABDITIDA: RHABDITIDIDEA: RHABDITIDAE: PELODERINHA: CAENORHABDITIS.				
AN	[1]				
RA	SEQUENCE FROM N.A.				
RA	BERKS M.;				
RT	SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.				
RP	[2]				
KP	SEQUENCE FROM N.A.				
EX	MEDLINE; 9416718.				
RA	WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,				
RA	BOUFIELD J., BURTON J., CONNELL M., COPESEY T., COOPER J., COULSON A.,				
RA	CRAXTON M., DEAR S., DU Z., DURBIN R., FAYELLO A., FULTON L.,				
RA	SARDEN A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,				
RA	JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LARREILLE P.,				
RA	LIGHTNING J., LLOYD C., MCMURRAY E., MORTIMORE B., O'CALLAGHAN M.,				
RA	PARSONS J., PERCY C., RIKKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,				
RA	SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,				
RA	THIERY-WIELE J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSOON R.,				
RA	WATSON A., WEINSTOCK L., WILKINSON-SPRAT J., WOHLIDAN P.,				
RT	3.2 Mb of contiguous nucleotide sequence from chromosome III of C.				
RT	elegans.				
FJ	NATURE 368:32-38(1994).				
DR	EMBL; Z70753; E1346477;				
SG	SEQUENCE 539 AA; 60947 MW; BDAE1002 CMC32;				

Query Match	11.48;	Score 125;	DB 5;	Length 539;
Best Local	Similarity 38.98;	Pred. No. 9.36e-02;		
Matches	21;	Conservative 18;	Mismatches 12;	Indels 3;
			Gaps 3;	
DB	193 FVCSICLNGTLLIYSLAVALPATSVERHLSVYCAIOTIALSISIVAFV 245			
QY	47 YI-SFGMSLGL-SLAVALTSPSVS-NALNWRREFSIQSLIGVALLSTFHV 97			
RESULT	2			

ID 050684 PRELIMINARY; PRT; 552 AA.

AC 050684;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)

DE PUTATIVE PHOSPHATE PERMEASE CY339.29C.

OS MTCY339.29C.

OS MYCOBACTERIUM TUBERCULOSIS.

OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;

OC ACTINOMYCETALES; CORYNEBACTERINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;

RA BADCOCK K., CHURCHER C.M., BARRELL B.G., RAJANDREAN M.A., WALSH S.V.;

RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- INNER MEMBRANE (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE PHO-4 FAMILY OF TRANSPORTERS.

DR EMBL: 277163; E255063.

KM HYPOTHETICAL PROTEIN; TRANSPORT; TRANSMEMBRANE; INNER MEMBRANE.

FT TRANSMEM 38 POTENTIAL.

FT TRANSMEM 69 POTENTIAL.

FT TRANSMEM 107 127 POTENTIAL.

FT TRANSMEM 146 166 POTENTIAL.

FT TRANSMEM 178 198 POTENTIAL.

FT TRANSMEM 213 233 POTENTIAL.

FT TRANSMEM 326 346 POTENTIAL.

FT TRANSMEM 360 380 POTENTIAL.

FT TRANSMEM 389 409 POTENTIAL.

FT TRANSMEM 437 457 POTENTIAL.

FT TRANSMEM 492 492 POTENTIAL.

FT TRANSMEM 493 513 POTENTIAL.

FT TRANSMEM 526 546 POTENTIAL.

SO SEQUENCE 552 AA; 58789 MW; 3164CDB8 CRC32;

Query Match 11.0%; Score 121; DB 2; Length 552;

Best Local Similarity 28.6%; Pred. No. 2,32e-01;

Matches 26; Conservative 21; Mismatches 39; Indels 5; Gaps 4;

DB 342 LKGFQKMLGLTTNNYFIAMGAAY-WMAFFIAKTLRG-ESLSRSTF--LMSMQ 397

QY 46 MYISFGIMSLGLSLAVTSIPSVSNALNMRE-FSFIQSLGYALLSTFHVLYIGMKR 104

DB 398 VFTASGFASHGSDINAIQFPAILDVLR 428

QY 105 AFEERYRYTPNPVLAIVLPISVILDLQ 135

FT 3 PRELIMINARY; PRT; 380 AA.

AC 063585;

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)

DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)

DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)

DE CYTOCHROME B.

OS FLOROMETRA SERRATISSIMA.

OC MITOCHONDRION.

OC EUKARYOTA; METAZOA; ECHINODERMATA; CRINOZOA; CRINOIDEA; ARTICULATA;

OC COMATULIDA; MACROPHREATA; ANTEDONIDAE; FLOROMETRA.

RN [1]

RP SEQUENCE FROM N.A.

RA SCOURAS A., SMITH M.J.;

RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q + 2

CC -1- FERROCYTOCHROME C.

CC COFACTOR: TWO HEME GROUPS

CC (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN

CC (BY SIMILARITY).

EMBL: AF049132; G2970421.

DR PROSTATE; PS00192; CYTOCHROME_B_HEME.1.

KM MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;

SO SEQUENCE 380 AA; 43086 MW; 69C0C55A CRC32;

Query Match 10.9%; Score 120; DB 8; Length 380;

Best Local Similarity 26.0%; Pred. No. 2,91e-01;

Matches 25; Conservative 31; Mismatches 32; Indels 8; Gaps 7;

DB 108 YVNEFMNNGVLLFLFYMAAFYGVLPWGMQSEFGATV-ITVLSAIPPLGYELVQWV 166

QY 47 YISFGIMSLGLSLAVTSIPSVSNALNMREFSFIQSLGYALLSTFHVLYIGMKR 104

DB 167 GGFVDNATLVRETF-HLLPFIITIALSVVLLFL 201

QY 105 A-F--EEERY-REYTPNPVLAIVLPISVILDLQ 136

RESULT 4

ID 021048 PRELIMINARY; PRT; 488 AA.

AC 021048;

DT 01-JAN-1998 (TREMBLREL. 05, CREATED)

DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE NADH DEHYDROGENASE SUBUNIT 2 (EC 1.6.5.3) (NADH DEHYDROGENASE

DE (UBIQUINONE)) (UBIQUINONE REDUCTASE) (TYPE I DEHYDROGENASE).

GN NAD2.

OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).

OC MITOCHONDRION.

OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-AX3;

RY MEDLINE: 97153471.

RA OGAWA S., MATSUO K., ANGATA K., YANAGISAWA K., TANAKA Y.;

RT "Group I" insertions in the cytochrome c oxidase genes of Dictyostelium

R1 discoideum: two related ORFs in one loop of a group-I intron, a

R2 cox1/2 hybrid gene and an unusually large cox3 gene.;"

CC CURR. GENE. 31:80-88(1997).

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.

CC -1- COFACTOR: FAD; IRON-SULFUR.

DR EMBL: D16579; D1021965;

DR PFAM: PF00361; oxidored_q1; 1.

KM OXIDOREDUCTASE; MITOCHONDRION.

SO SEQUENCE 488 AA; 54841 MW; 4A899DB7 CRC32;

Query Match 10.6%; Score 116; DB 8; Length 488;

Best Local Similarity 27.1%; Pred. No. 7,08e-01;

Matches 23; Conservative 31; Mismatches 27; Indels 4; Gaps 4;

DB 268 YLNEVINKGLVLLVLTGVTGSL-ROQ-KVIRF-INYSAIVNSALLLFEVGNNT 324

QY 47 YISFGIMSLGLSLAVTSIPSVSNALNMREFSFIQSLGYALLSTFHVLYIGMKR 106

DB 325 ELIYSIYLYINIIIGLAVLNIIT 349

QY 107 EEEYRYTPNPVLAIVLPISVIL 130

RESULT 5

ID 013397 PRELIMINARY; PRT; 1055 AA.

AC 013397;

DT 01-JAN-1998 (TREMBLREL. 05, CREATED)

DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE P-TYPE ATPASE 1.

GN ENA1.

OS DEBARYOMYCES OCCIDENTALIS.

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;

OC SACCCHAROMYCETACEAE; DEBARYOMYCES.

RN [1]

RP SEQUENCE FROM N.A.

RA BANDLOS M.A., RODRIGUEZ-NAVARRO A.;

RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF030860; G2623236;

DR PFAM: PF00122; E1-E2_ATPase; 2.

SQ SEQUENCE 1055 AA; 116651 MW; 5A9F92CE CRC32;
 Query Match 10.6%; Score 116; DB 3; Length 1055;
 Best Local Similarity 24.0%; Pred. No. 7.08e-01;
 Matches 23; Conservative 37; Mismatches 26; Indels 10; Gaps 10;

Db 934 LRSFFMRPELEPMWKKOALIDMDQ-FLFVSMG-AIVSVPVYIPIVINKVELHA 991
 21 LMAAYQGVHNIENSM-NE-E-EVWRIEMYSFGIMSLGLSLAVSIPSVNALMRE 77
 Y 78 -FSFIQSTLGYALLISTFHVLIYGMKRAFEVEEYR 112

RESULT 6 PRELIMINARY; PRT: 472 AA.
 AC 94422;
 DT -MAY-1997 (TREMBLREL. 03, CREATED)
 DT -MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
 DT -NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HOMOLOGUE OF MULTIDRUG RESISTANCE PROTEIN B, EMR8, OF E. COLI.
 GN YCNB.
 OC BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 OC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 TRPC2;
 RX MEDLINE: 97124189.
 RA YAMANE K., KUMANO M., KURITA K.;
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis
 chromosome: determination of the sequence of a 146 kb segment and
 identification of 113 genes.";
 RL MICROBIOLOGI 142:3047-3056(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE: 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
 BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
 BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 DENOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
 ENTIAN K.D., ERINGTON J., FABRET C., FERRARI E., FOULGER D.,
 GARTY C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,
 GILBERT S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
 GUISSEPI G., GUY B.J., HAGA K., HATECH J., HARWOOD C.R., HENAUT A.,
 HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
 JORIS B., KARASAKI D., KASAHARA Y., KLAER-BLANCARD M., KLEIN C.,
 KOBAYASHI Y., KOETTER P., KONINGSSTEIN G., KROGH S., KUMANO M.,
 KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 NOONE D., O'REILLY M., OGAMA K., OGIMARA A., OUDGA B., PARK S.H.,
 PARRO V., POHL T.M., PORTELELLA D., PORTELLI S., PRESCOTT A.M.,
 PRESCAN E., PUJIC P., PYRELLIE B., RAPOPORT G., REY M., REYNOLDS S.,
 RIEGER M., RIVOLTA C., ROCHA E., ROCHÉ B., ROSE M., SAADE Y.,
 SATO T., SCHNLAN E., SCHLECH S., SCHROETER R., SCOFONE F.,
 SEKICUCHI J., SEKOWSKA A., SEROR S.J., SEROR P., SHIN B.S., SOLDO B.,
 SOROKIN A., TACCONE E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 TAKEUCHI M., TAKAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,
 TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
 VIARI A., WAMBOIT R., WEDER E., WEDLER H., WEITENEGGER T.,
 WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
 YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 subtilis.";
 RL NATURE 390:249-256(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;

RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN EMBL: D50453; G1805454;
 RN EMBL: 295106; E1182351;
 SQ SEQUENCE 472 AA; 50797 MW; 2836006E CRC32;

Query Match 10.5%; Score 115; DB 2; Length 472;
 Best Local Similarity 27.1%; Pred. No. 8.83e-01;
 Matches 23; Conservative 37; Mismatches 34; Indels 5; Gaps 4;

Db 145 LVSFAPAIPTYSNAVAEAFSWRSLEIILPFAVIDLIASLI-MKNVTLTKTOI-DI 202
 21 LMAAYQGVHNIENSMNEE-VRIEMXI--SPGIMSLGLSLAVSIPSVNALMRE 77
 Y 78 -FSFIQSTLGYALLISTFHVLIYGM 102

RESULT 7 PRELIMINARY; PRT: 256 AA.
 AC 04638;
 DT -JAN-1998 (TREMBLREL. 05, CREATED)
 DT -JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT -NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 29.9 KD PROTEIN.
 GN SAIMIRINE HERPESVIRUS 2.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
 OC GAMMAHERPESVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-488;
 RX MEDLINE: 98037620.
 RA KNAPE A., HILDER C., THURAU M., WITTMANN S., HOFMANN H.,
 RA FLECKENSTEIN B., FICKENSCHER H.;
 RT "The superantigen-homologous viral immediate-early gene iel4/vsag in
 herpesvirus saimiri-transformed human T cells.";
 RL J. VIROL. 71:9124-9133(1997).
 DR EMBL: Y13183; E332528;
 RT HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 256 AA; 29889 MW; 5720E5A2 CRC32;

Query Match 10.2%; Score 112; DB 14; Length 256;
 Best Local Similarity 25.7%; Pred. No. 1.70e+00;
 Matches 27; Conservative 31; Mismatches 40; Indels 7; Gaps 7;

Db 101 IDPS-SRGQMLHNSNDILQGMQSFISKLQIHFVRISPNITDGNLFGVLPY 159
 32 IENSMNEEYV-RLEMYISF-GIMSLGLSLAVTSIPSVS-NALMNRHSF-IQSTLGY 87
 Y 88 VALLISTFHVLIYGMKRAFEVEEYRPTPNFVALVLPISVILD 132

RESULT 8 PRELIMINARY; PRT: 386 AA.
 AC 06545;
 DT -AUG-1998 (TREMBLREL. 07, CREATED)
 DT -AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT -NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 43.8 KD PROTEIN.
 GN A0155.
 OS AQUIFEX AEOIICUS.
 OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBERG R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON S., SWANSON R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex

```

RT aeolicus."
RL NATURE 392:353-358(1998).
RN
RN [2]
RN SEQUENCE FROM N.A.
RC STERAIN-VF5;
RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG M.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJAYI M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.,
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AE00676; G2982892;
KW HYPOTHEICAL PROTEIN.
SQ SEQUENCE 386 AA: 43793 MW: A4F7E62 CRC32;

Query Match 10.0%; Score 112; DB 2; Length 386;
Best Local Similarity 25.9%; Pred. No. 1.70e+00;
Matches 21; Conservative 27; Mismatches 25; Indels 8; Gaps 7;

DB 2 LW-FSLFI-LSIAGSFAIIVARTPGADLFP-PKY-FYALVGHVDSALIVGLYAF 57
QY 41 VRIEIKYISFGISLGLSL-L-AVTSIPVSALNMRREFSFIQSTIGYV--ALLISTFHY 97
DB 58 LIFLWHRIFEKKENFASFLPA 78
QY 98 LLYGMKRAFEF-EYRYFTPP 117

RESULT 9
ID 019916 PRELIMINARY; PRT; 205 AA.
AC 019916;
DT 01-JAN-1998 (TREMREL. 05, CREATED)
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE YCF15 PROTEIN.
GN YCF15.
OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OC EUKARYOTA.
OC EUKARYOTA: RHODOPHYTA: BANGIOPHYCEAE: PORPHYRIDIALES: PORPHYRIDACEAE;
OC CYANIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RK1;
RA GLOECKNER G., ROSENTHAL A., VALENTIN K.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF02186; G2465744;
KW CHLOROPLAST.
SQ SEQUENCE 205 AA: 23273 MW: FBE22A8F CRC32;

Query Match 10.0%; Score 110; DB 8; Length 205;
Best Local Similarity 34.8%; Pred. No. 2.61e+00;
Matches 24; Conservative 13; Mismatches 29; Indels 3; Gaps 3;

DB 121 QISKFTKFIHVALLLYLD-SVIGTVLTSPLSIGSYFIDRLADTLILI-TFMISIVA 178
QY 43 RIEMYSFGIMSLGLSLAVTSIPVSALNMR-EFSFIQSTIGYVALLISTFHYLYIG 101
DB 179 GTDALGOY 187
QY 102 WKRAFEERY 110

RESULT 10
ID P90817 PRELIMINARY; PRT; 319 AA.
AC P90817;
DT 01-MAY-1997 (TREMREL. 03, CREATED)
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)
DE F09F3.1 PROTEIN.
GN CAENORHABDITIS ELEGANS.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA: METAZOA: NEMATODA: SECERNITEA: RHABDITIA: RHABDITIDA;
OC RHABDITIDA: RHABDITOIDEA: RHABDITIDAE: PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA THOMAS K.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,
RA CRATON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES N., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SUSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL NATURE 368:32-38(1994).
DR EMBL: Z81056; E1345334;
SQ SEQUENCE 319 AA: 36986 MW: BD9AD661 CRC32;

Query Match 10.0%; Score 110; DB 5; Length 319;
Best Local Similarity 24.7%; Pred. No. 2.61e+00;
Matches 24; Conservative 26; Mismatches 40; Indels 7; Gaps 7;

DB 197 ITNFENIAIV-ITLMSMSGVNAESRRARRKRW-NFTQCVIQDSLOLFTINAY-IM 253
QY 44 IEMVIFGIMSLGLSL-L-A-VTSIPVSALNMRREFSFIQSTIGYVALLISTFHYLYIG 101
DB 254 W-RFYDAWYRFLEATLSILVSLDGFVMTYHOEC 289
QY 102 WKRAFEERYRF-YTPNFVALVLPSTIYIIDLQIC 137

RESULT 11
ID P87410 PRELIMINARY; PRT; 381 AA.
AC P87410;
DT 01-MAY-1997 (TREMREL. 03, CREATED)
DT 01-MAY-1997 (TREMREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE CYTOCHROME B.
GN CYTB.
OS MYRMECOBIUS FASCIAIUS.
OS MITOCHONDRION.
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: METATHERIA;
OC DASYROMORPHIA: MYRMECOBIIDAE: MYRMECOBIUS.
RN [1]
RP SEQUENCE FROM N.A.
RA KRAJEWSKI C., PAINTER J., DRISKELL A.C., BUCKLEY L., WESTERMAN M.;
RI SCI. NEW GUINEA 19:157-166(1993).
RN [2]
RN SEQUENCE FROM N.A.
RA KRAJEWSKI C., WESTERMAN M.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C -> Q + 2
CC FERROCYTOCHROME C.
CC -1- COFACTOR: TWO HEME GROUPS
CC (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN
CC (BY SIMILARITY).
DR EMBL: U82329; G1785600;
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PFAM: PF00032; cytochrome_b_c; 1.
DR PFAM: PF00033; cytochrome_b_n; 1.
KW MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
KW HEME.
SQ SEQUENCE 381 AA: 42795 MW: DA0B368 CRC32;

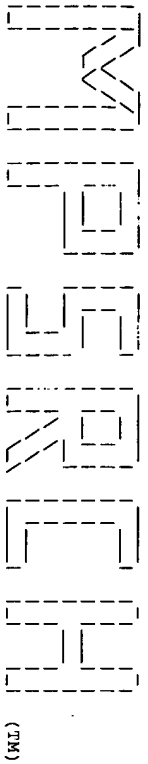
Query Match 9.9%; Score 109; DB 8; Length 381;
Best Local Similarity 25.8%; Pred. No. 3.24e+00;
Matches 23; Conservative 28; Mismatches 30; Indels 8; Gaps 8;

DB 114 NIGVILLTVMTAFVGYLPMGQMSFMGATV-ITLLSAIFYIGSLYEWIMGGSVDK 172

```


(BY SIMILARITY).
CC EMBL: AF020234; G2444365; -
DR EMBL: AF020253; G2444363; -
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PFAM: PF00032; cytochrome_b_C; 1.
DR PFAM: PF00033; cytochrome_b_N; 1.
KM MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
KW HEME.
FT NON_TER 1 1
FT NON_TER 300 300
SQ SEQUENCE 300 AA; 33590 MW; CBC69B4E CRC32;
Query Match 9.88; Score 108; DB 8; Length 300;
Best Local Similarity 29.28; Pred. No. 4.01e+00;
Matches 26; Conservative 24; Mismatches 31; Indels 8; Gaps 7;
Db 97 NLGVILLVMTAFVGYVLPWGMSPFGATV-ITNLLSATPYVGTLLVEMWGFSDN 155
QY 54 SLGLSLAVTSIPSYSNALNMRFSFIOSTLGVALLIST-FH-VLIYGMKRA-F--EE 108
D 156 ALTLREFTF-HLLPFVIIGAAVLAHLLEL 183
QY 109 EYX-REYTPPNFVLAIVLPSIVILDLQL 136

Search completed: Thu Jan 27 17:34:20 2000
Job time : 21 secs.



Release 3.1A John F. Collins, B.Occupational Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Message: protein - protein database search, using Smith-Waterman algorithm
Run: Thu Jan 27 17:40:11 2000; Maspar time 7.52 Seconds
700.292 Million cell updates/sec
Tabular output not generated.

Title: >US-09-397-558-2
Description: (1-410) from US09397558.pep
Perfect Score: 3201
Sequence: 1 MEPLPVVLAIKRSRYVLEAAV.....RDCRMGAGTGVAVGILV 410

Scoring table: PAM 150
Gap 11

Searched: 130303 seqs, 12848679 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfile1

Statistics: Mean 33.711; Variance 155.316; scale 0.217

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Re No.	Score	Query Match	Length	ID	Description	Pred. No.
1	781	24.4	229	1	US-08-702-Sequence 9, Applicatio	2.97e-58
2	123	3.8	334	1	US-08-559-Sequence 2, Applicatio	3.33e-01
3	110	3.4	591	1	US-08-483-Sequence 2, Applicatio	2.86e+00
4	110	3.4	591	1	US-08-484-Sequence 2, Applicatio	2.86e+00
5	101	3.2	265	3	PCT-US95-0-Sequence 29, Applicatio	1.21e+01
6	101	3.2	265	1	US-08-413-Sequence 27, Applicatio	1.21e+01
7	101	3.2	286	1	US-08-321-Sequence 5, Applicatio	1.03e+01
8	102	3.2	362	2	US-08-802-Sequence 5, Applicatio	1.03e+01
9	102	3.2	1618	1	US-07-853-Sequence 2, Applicatio	1.03e+01
10	102	3.2	1805	1	US-07-853-Sequence 2, Applicatio	1.03e+01
11	98	3.1	1452	2	US-08-087-Sequence 8, Applicatio	1.94e+01
12	98	3.1	1452	2	US-08-449-Sequence 8, Applicatio	1.94e+01
13	95	3.0	1452	2	US-08-991-Sequence 4, Applicatio	3.09e+01
14	95	3.0	1452	2	US-08-652-Sequence 4, Applicatio	3.09e+01
15	95	3.0	1452	2	US-08-769-Sequence 4, Applicatio	3.09e+01
16	94	2.9	98	3	PCT-US93-0-Sequence 336, Applicat	3.60e+01
17	94	2.9	98	1	US-08-118-Sequence 336, Applicat	3.60e+01
18	92	2.9	316	1	US-08-414-Sequence 22, Applicati	4.89e+01
19	92	2.9	316	2	US-08-926-Sequence 22, Applicati	4.89e+01
20	92	2.9	317	2	US-08-790-Sequence 3, Applicatio	4.89e+01
21	92	2.9	317	2	US-09-213-Sequence 3, Applicatio	4.89e+01
22	92	2.9	317	2	US-08-619-Sequence 8, Applicatio	4.89e+01
23	92	2.9	340	2	US-09-213-Sequence 1, Applicatio	4.89e+01

24	92	2.9	340	2	US-08-790-Sequence 1, Applicatio	4.89e+01
25	93	2.9	361	1	US-08-415-Sequence 7, Applicatio	4.20e+01
26	94	2.9	593	1	US-08-296-Sequence 2, Applicatio	3.60e+01
27	53	2.9	890	2	US-08-451-Sequence 8, Applicatio	4.20e+01
28	93	2.9	890	1	US-08-472-Sequence 8, Applicatio	4.20e+01
29	93	2.9	890	2	US-08-472-Sequence 8, Applicatio	4.20e+01
30	93	2.9	890	2	US-08-472-Sequence 8, Applicatio	4.20e+01
31	90	2.8	96	3	PCT-US93-0-Sequence 340, Applicat	6.63e+01
32	90	2.8	96	3	PCT-US93-0-Sequence 340, Applicat	6.63e+01
33	90	2.8	98	3	PCT-US93-0-Sequence 337, Applicat	7.71e+01
34	90	2.8	342	2	US-08-483-Sequence 2, Applicatio	6.63e+01
35	90	2.8	370	2	US-08-031-Sequence 2, Applicatio	6.63e+01
36	90	2.8	384	1	US-08-232-Sequence 4, Applicatio	6.63e+01
37	90	2.8	384	3	PCT-US93-0-Sequence 3, Applicatio	6.63e+01
38	91	2.8	470	2	US-08-555-Sequence 15, Applicati	6.63e+01
39	89	2.8	1107	3	PCT-US96-0-Sequence 1, Applicatio	5.70e+01
40	89	2.8	1107	1	US-08-366-Sequence 2, Applicatio	7.71e+01
41	89	2.8	1229	1	US-08-176-Sequence 4, Applicatio	7.71e+01
42	89	2.8	1229	1	US-08-779-Sequence 4, Applicatio	7.71e+01
43	89	2.8	1436	2	US-08-652-Sequence 2, Applicatio	7.71e+01
44	89	2.8	1436	2	US-08-991-Sequence 2, Applicatio	7.71e+01
45	89	2.8	2183	1	US-08-348-Sequence 7, Applicatio	7.71e+01

ALIGNMENTS

RESULT 1
US-08-702-344-9 STANDARD: PRT: 229 AA.
xxxxxx
Sequence 5: Application US/08702344
Sequence 9: Application US/08702344
Patent No. 5723315
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Lavalley, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version v1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,344
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 229 amino acids


```

ID      US-08-484-840-2    STANDARD:    PRT:    591 AA.
XX      xxxxxx
DT
Sequence 2, Application US/08484840
CC      Sequence 2, Application US/08484840
CC      Patent No. 5716788
CC      GENERAL INFORMATION:
CC      APPLICANT: MOSCOW, Jeffrey A.
CC      APPLICANT: COMAN, Kenneth H.
CC      APPLICANT: DIXON, Kathy
CC      APPLICANT: HE, Rui
CC      TITLE OF INVENTION: A GENE ENCODING A HUMAN REDUCED FOLATE
CC      TITLE OF INVENTION: CARRIER (RFC) AND METHODS FOR THE TREATMENT OF
CC      TITLE OF INVENTION: METHOTREXATE-RESISTANT, TRANSPORT-DEFICIENT CANCER CELLS
CC      NUMBER OF SEQUENCES: 6
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Foley & Lardner
CC      STREET: 3000 K Street, N.W., Suite 500
CC      CITY: Washington
CC      STATE: D.C.
CC      COUNTRY: USA
CC      ZIP: 20007-5109
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent In Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/484,840
CC      FILING DATE: 07-JUN-1995
CC      CLASSIFICATION: 530
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: BENT, Stephen A.
CC      REGISTRATION NUMBER: 29,768
CC      REFERENCE/DOCKET NUMBER: 40399/324/NHHD
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202)672-5300
CC      TELEFAX: (202)672-5399
CC      TELEX: 904136
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 591 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ      SEQUENCE 591 AA; 64921 MW; 1822494 CN;
Query Match          3.4%; Score 110; DB 1; Length 591;
Best Local Similarity 31.8%; Pred. No. 2,86e+00;
Matches 28; Conservative 22; Mismatches 29; Indels 9; Gaps 9;
Db      192 FLTFSVVALFLKRPRLSFF-NRDDRGCESTASSELERNPNPGGKIGALRVACG-DS 249
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      243 FLEFFGVSPILRR-RQAVFLIEDMESFRTONSSRDLEPF-PGHGE-LPEGLSPCIMES 299
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      250 VLARMLRELG-DSL-R-PQLRLWSLW 274
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      300 FL-RTGAAYAGTESLRTRKESLIQYWSLSW 326
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT      5
ID      PCT-US95-03776-29    STANDARD:    PRT:    265 AA.
XX      xxxxxx
XX

```

[illegible]

CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0345 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 362 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GenBank
CC CLONE: 1123105
SQ SEQUENCE 362 AA; 43020 MW; 726128 CN;

Query Match 3.2%; Score 102; DB 2; Length 362;
Best Local Similarity 23.1%; Pred. No. 1.03e+01;
Matches 9; Conservative 19; Mismatches 11; Indels 0; Gaps 0;

Db 151 SLIGSTFVDRSDFWOLMVHVTITFLSSMTINPVR 189
QY 85 AMLSMALQLDRLHGLWMLGSPFLGLTATMTVRSYR 123

RESULT 9
ID US-07-853-913-4 STANDARD; PRT; 1618 AA.
XX
AC xxxxxx
DT
DE Sequence 4, Application US/07853913
CC Sequence 4, Application US/07853913
CC Patent No. 5338839
CC GENERAL INFORMATION:
CC APPLICANT: McKay, Ronald D.G.
CC APPLICANT: Lendahl, Urban
CC TITLE OF INVENTION: Nestin Expression As An Indicator of
CC TITLE OF INVENTION: Neuroepithelial Tumors
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/853,913
CC FILING DATE: 19920319
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/660,412
CC FILING DATE: 22-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/603,803
CC FILING DATE: 25-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/201,762
CC FILING DATE: 02-JUN-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/180,548
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227

CC REFERENCE/DOCKET NUMBER: MIT-4641AAAA
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1618 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1618 AA; 176705 MW; 12580209 CN;

Query Match 3.2%; Score 102; DB 1; Length 1618;
Best Local Similarity 23.0%; Pred. No. 1.03e+01;
Matches 20; Conservative 23; Mismatches 43; Indels 1; Gaps 1;

Db 138 EALRVAHEERY-GLNQAACAPRLPAPPPAPPEVEELARLGAMGAVRGIOERY 196
QY 310 ESLTKESLLQVMSLSWDAPSDMDSPGRQSPVRSSTASFORWISLWGNQISRSQRL 369
Db 197 AHMTSLDQTRERLARAVOGAREVYLE 223
QY 370 SNGSLRIPSGORGLGCAVLMRDCRMD 396

RESULT 10
ID US-07-853-913-2 STANDARD; PRT; 1805 AA.
XX
AC xxxxxx
DT
DE Sequence 2, Application US/07853913
CC Sequence 2, Application US/07853913
CC Patent No. 5338839
CC GENERAL INFORMATION:
CC APPLICANT: McKay, Ronald D.G.
CC APPLICANT: Lendahl, Urban
CC TITLE OF INVENTION: Nestin Expression As An Indicator of
CC TITLE OF INVENTION: Neuroepithelial Tumors
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/853,913
CC FILING DATE: 19920319
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/660,412
CC FILING DATE: 22-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/603,803
CC FILING DATE: 25-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/201,762
CC FILING DATE: 02-JUN-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/180,548
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: MIT-4641AAAA

```
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617-861-6240  
CC TELEFAX: 617-861-9540  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1805 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 1805 AA; 198744 MW; 15616656 CN;  
  
Query Match 3.2%; Score 102; DB 1; Length 1805;  
Best Local Similarity 26.7%; Pred. No. 1.03e+01;  
Matches 16; Conservative 14; Mismatches 29; Indels 1; Gaps 1  
  
Db 166 PHRTGPAPVEYEDLARRIGEVWRCGVARDYQDERVAMESLSGQARPRLSQAVRGARECTLE 225  
Oy 338 PGR-QSPVRSTASFQRKRWLSWGNSISRFSSQLSNSGLRPLSQORRLCGAVALMRDRCDMD 396  
R 11  
ID US-08-087-244A-8 STANDARD; PRT; 1452 AA.  
XX xxxxxx  
DT  
XX  
DE Sequence 8, Application US/08087244A  
XX Sequence 8, Application US/08087244A  
CC Patent No. 5863755  
CC GENERAL INFORMATION:  
CC APPLICANT: Schlessinger, Joseph  
CC APPLICANT: Sap, Jan M.  
CC APPLICANT: Ullrich, Axel  
CC APPLICANT: Vogel, Wolfgang  
CC APPLICANT: Fuchs, Miriam  
CC TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE  
CC TITLE OF INVENTION: PHOSPHATASE-KARPA  
CC NUMBER OF SEQUENCES: 11  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: PENNIE & EDMONDS  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/087,244A  
CC FILING DATE: 01-JUL-1993  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Coruzzi, Laura A.  
CC REGISTRATION NUMBER: 30,742  
CC REFERENCE/DOCKET NUMBER: 7683-042  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 212-790-9090  
CC TELEFAX: 212-869-8664/9741  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1452 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 1452 AA; 163632 MW; 11201300 CN;
```

Query Match	3.1%	Score 98	DB 2	Length 1452
Best Local Similarity	18.3%	Pred. No. 1.94e+01		
Matches	20	Conservative	37	Mismatches 48; Indels 4; Gaps 4;

Db	887	QHTOMKCAEYGEKREYESFEEOQSAKPDMSAKDENMKNRGNITAYDHSRYRLQTE	946
Qy	224	QHLCOLRGLALARG-YECLEFFELVWSPFLRRQAVLEDMESFSRTONSSNDLEFPF	282
Db	947	GDTN-SDVINGNYI-DGYHRPNHIATOG-PMOETIYDFNRVMWHENTA	992
Qy	283	GHGELPELESPCIMESFLRTGAVAGTIESLRTKESLLQVMSLSDAEP	331

RESULT	12	STANDARD	PRT	1452 AA.
ID	US-08-449-644-8			
AC	xxxxxx			
XX				
D7				
XX				
DE				
XX				
Sequence 8, Application US/08449644				
Sequence 8, Application US/08449644				
Patent No. 5856162				
GENERAL INFORMATION:				
APPLICANT: Schliesinger, Joseph				
APPLICANT: Sap, Jan M.				
APPLICANT: Ullrich, Axel				
APPLICANT: Vogel, Wolfgang				
APPLICANT: Fuchs, Miriam				
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE				
TITLE OF INVENTION: PHOSPHATASE-KAPPA				
NUMBER OF SEQUENCES: 11				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: PENNIE & EDMONDS				
STREET: 1155 Avenue of the Americas				
CITY: New York				
STATE: New York				
COUNTRY: U.S.A.				
ZIP: 10036				
COMPUTER READABLE FORM:				
MEDIUM TYPE: Floppy disk				
COMPUTER: IBM PC compatible				
OPERATING SYSTEM: PC-DOS/MS-DOS				
SOFTWARE: Patentin Release #1.0, Version #1.25				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/08/449,644				
FILING DATE: 24-MAY-1995				
CLASSIFICATION: 514				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: US 08/087,244				
FILING DATE: 01-JUL-1993				
ATTORNEY/AGENT INFORMATION:				
NAME: Coruzzi, Laura A.				
REGISTRATION NUMBER: 30,742				
REFERENCE/DOCKET NUMBER: 7683-042				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: 212-790-9090				
TELEFAX: 212-869-8864/9741				
TELEX: 66141 PENNIE				
INFORMATION FOR SEQ ID NO: 8:				
SEQUENCE CHARACTERISTICS:				
LENGTH: 1452 amino acids				
TYPE: amino acid				
STRANDEDNESS: single				
TOPOLOGY: unknown				
MOLECULE TYPE: protein				
SEQUENCE	1452 AA; 163632 MW; 11201300 CN;			

Query Match	3.1%	Score 98	DB 2	Length 1452
Best Local Similarity	18.3%	Pred. No. 1.94e+01		
Matches	20	Conservative	37	Mismatches 48; Indels 4; Gaps 4;

Db 887 QHTOMKACBEGYKFEYESFEGOSAPWDSAKDENMKRNGNIIAYDSRVYLQIE 946
 QY 224 QHLCOLRGALGLARG-YECFLFEFLGWSPRLRRQAVFLDEMSFSTQSSRDLEFPF 282
 Db 947 GDTN-SOYINGNYI-DGYHRPNHYIATOG-PMOETIYDFWRMWHENTA 992
 QY 283 GHGELPGLESPICMESFLRTGAYAGTESLRTKESLLQVMSLSWDAEPS 331

RESULT 13
 ID US-08-991-258A-4 STANDARD: PRT: 1452 AA.

XX AC xxxxxx
 CC Sequence 4, Application US/08991258A
 CC Patent No. 5928887
 CC GENERAL INFORMATION:
 CC APPLICANT: Cheng, Jili
 CC APPLICANT: Lasky, Laurence A.
 CC TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
 CC NUMBER OF SEQUENCES: 10
 CC CORRESPONDENCE ADDRESSES:
 CC ADDRESS: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP
 CC STREET: 4 Embarcadero Center, Suite 3400
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: United States
 CC ZIP: 94111
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/991,258A
 CC FILING DATE: 17-DEC-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/652,971
 CC FILING DATE: 24-MAY-1996
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Dreger, Walter H.
 CC REGISTRATION NUMBER: 24,190
 CC REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTK
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 781-1989
 CC TELEFAX: (415) 398-3249
 CC TELEX:
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1452 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 1452 AA; 163593 MW; 11130193 CN;
 CC
 CC Query Match 3.0%; Score 95; DB 2; Length 1452;
 CC Best Local Similarity 18.3%; Pred. No. 3.09e+01;
 CC Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;
 Db 887 QHTOMKACBEGYKFEYESFEGOSAPWDSAKDENMKRNGNIIAYDSRVYLQIE 946
 QY 224 QHLCOLRGALGLARG-YECFLFEFLGWSPRLRRQAVFLDEMSFSTQSSRDLEFPF 282
 Db 947 GDTN-SOYINGNYI-DGYHRPNHYIATOG-PMOETIYDFWRMWHENTA 992
 QY 283 GHGELPGLESPICMESFLRTGAYAGTESLRTKESLLQVMSLSWDAEPS 331

RESULT 14
 ID US-08-652-971-4 STANDARD: PRT: 1452 AA.

XX AC xxxxxx
 CC Sequence 4, Application US/08652971
 CC Patent No. 5814507
 CC GENERAL INFORMATION:
 CC APPLICANT: Cheng, Jili
 CC APPLICANT: Lasky, Laurence A.
 CC TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
 CC NUMBER OF SEQUENCES: 10
 CC CORRESPONDENCE ADDRESSES:
 CC ADDRESS: Genentech, Inc.
 CC STREET: 460 Point San Bruno Blvd.
 CC CITY: South San Francisco
 CC STATE: California
 CC COUNTRY: United States
 CC ZIP: 94080
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/652,971
 CC FILING DATE:
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Dreger, Ginger R.
 CC REGISTRATION NUMBER: 33,055
 CC REFERENCE/DOCKET NUMBER: P1033
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 225-3216
 CC TELEFAX: (415) 952-9881
 CC TELEX: 910 371-7168
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1452 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 1452 AA; 163593 MW; 11130193 CN;
 CC
 CC Query Match 3.0%; Score 95; DB 2; Length 1452;
 CC Best Local Similarity 18.3%; Pred. No. 3.09e+01;
 CC Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;
 Db 887 QHTOMKACBEGYKFEYESFEGOSAPWDSAKDENMKRNGNIIAYDSRVYLQIE 946
 QY 224 QHLCOLRGALGLARG-YECFLFEFLGWSPRLRRQAVFLDEMSFSTQSSRDLEFPF 282
 Db 947 GDTN-SOYINGNYI-DGYHRPNHYIATOG-PMOETIYDFWRMWHENTA 992
 QY 283 GHGELPGLESPICMESFLRTGAYAGTESLRTKESLLQVMSLSWDAEPS 331
 RESULT 15
 ID US-08-769-399-4 STANDARD: PRT: 1452 AA.
 XX AC xxxxxx
 CC Sequence 4, Application US/08759399

XX Sequence 4, Application US/08769399
CC Patent No. 5976852
CC GENERAL INFORMATION:
CC APPLICANT: Cheng, Jill
CC APPLICANT: Lasky, Laurence A.
CC TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
CC TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
CC NUMBER OF SEQUENCES: 10
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genentech, Inc.
CC STREET: 460 Point San Bruno Blvd.
CC CITY: South San Francisco
CC STATE: California
CC COUNTRY: United States
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/769,399
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dreger, Ginger R.
CC REGISTRATION NUMBER: 33,055
CC REFERENCE/DOCKET NUMBER: P1033
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 225-3216
CC TELEFAX: (415) 952-9881
CC TELEX: 910 371-7168
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1452 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1452 AA; 163593 MW; 11130193 CN;
SQ
Query Match 3.0%; Score 95; DB 2; Length 1452;
Best Local Similarity 18.38; Pred. No. 3.09e+01;
Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;
Db 887 QHITQKCAEGYGFKEEYESFFEGOSAPMDSAKKDEMRKNGITAYDHSRVRLQMLE 946
Q 224 QHICQLRGALGLALRG-YECFLFEFLGVWSPFLRRQAVFLDEMSFSRTONSSRDLEFPF 282
947 GDNN-SDYINGNYI-DGYHRPNHYIATQG-PMQETIYDFWRVWMENTA 992
OY 283 GHGELPEGLSEPCIMESFLRTGAVAGTESLRTKESLLQVWSLSWDAPPS 331

Search completed: Thu Jan 27 17:40:29 2000
Job time : 18 secs.

Query Match	3.6%	Score 115;	DB 8;	Length 367;
Best Local Similarity	24.1%;	Pred. No. 6.60e-02;		
Matches	28;	Conservative	31;	Mismatches 56;
			Indels	1;
			Gaps	1;

[illegible]

RESULT	6	PRELIMINARY:	PR:	405 AA.
ID	083354			
AC	083354:			
DT	01-NOV-1998	(TREMBLREL. 08. CREATED)		
	01-NOV-1998	(TREMBLREL. 08. LAST SEQUENCE UPDATE)		
	01-NOV-1998	(TREMBLREL. 08. LAST ANNOTATION UPDATE)		
	CONSERVED	HYPOTHETICAL PROTEIN.		

OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; TREPONEMA.

RP SEQUENCE FROM N.A.
RX MEDLINE: 98332770.

RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM
BA SCHREIBER F., HARRIS T.M., WATSON W.C., CARRICO C.

RÀ
MCDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII
KÀ KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M.

RA VENTER J.C.;
RT "Complete Genome Sequence of *Treponema pallidum*, the

SCIENCE 261:375-388(1998).

RP SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., S

RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S.,
RA KHALIL H., RICHARDSON D., HOWELL T.R., CHIDAMBARAM M.

RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J.

RECEIVED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL: AE001213; G33322611; -

Quercus	Match	360	Score	115	RD	3	1000	1000
Quercus	Match	360	Score	115	RD	3	1000	1000

Best Local Similarity	35.58;	Pred. No. 6.60e-02;	
Matches	22; Conservative	14; Mismatches	22; Ind

Db 93 WGSFPL-LKGVACVLFGLGVLYYA-RHRALGFLSRIVFFGQAQPTP

230 KQALSLPAAVLEIC - I DEL I QVWSP LAKKQAVE IEDWESF SKI QNNS

QY 288 PE 289

DESIGN 2

ID	080823	PRELIMINARY;	PRI;	24 / AA.
AC	086825;			

DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)

SC8B7.04.

OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDA

CC	MALINOXCELLIDES; SINEPIOMILINAE; SINEPIOMICELACAE; SINEPIOMICES
RN	[1]
RT	SEQUENCE FROM N.A.
EC	STRAIN-A3(2);

RL, *SUBMITTED (AUG-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
 RM [2]
 RN
 RP SEQUENCE FROM N.A.
 RQ STRAIN-A3(2).
 RA PARHILL J., BARRELL B.G., RAJANDEAN M.A.,
 RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
 RM [3]

RX MEDLINE: 97000351.
 RA REDENBACH M., KIESER H.M., DENNAPATE D., EICHNER A., CULLUM J.,
 KINSHI H., HOPWOOD D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map
 of the 8 Mb streptomyces coelicolor A3(2) chromosome.";
 FT MOL. MICROBIOL., 21:77-96(1996).
 RI EMBL: AL031225; E1314651; -
 RZ

Query Match 3.5%; Score 111; DB 2; Length 247;

Matches 24; Conservative 14; Mismatches 22; Indels 4; Gaps

[illegible]

215 VERB 218

QY 122 VRRR 125

RESULT 8

AC, 034431;
DT 01-JAN-1998 (TREMBLREL, 05, CREATED)

01-NOV-1998 (TREMBLEL. 08, LAST ANNOTATION UPDATE)

GN YLOB.
OS BACILLUS SUBTILIS.

OC PACILLUS.
RN [1]

AT SEQUENCE FROM N.A.
RC STRAIN=168;
.....

RA. KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA. AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,

RA. BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M., CHOI S.Y., CODANT T., COURENCHON T.D., CHIRIVICCA N.T., DAUTER D.

FA. DENIZOT F., DEVINE K.M., DOSTEROFF A., EHRlich S.D., EMERSON P.T.
ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,

KA CHIM-S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
RA GIUSEPPI G., GUY B.J., HAGA K., HATECH I., HARWOOD C.B., HENAUIT A.

JORIS B., KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,

RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,

RA NOONE D., O'REILLY M., OGIMARA A., OUDEGA B., PARK S.H.,
RA PABRO V. BOHLE T M BOETTCHER D BOHLE T M BOETTCHER S
BOHLE T M BOETTCHER S BOHLE T M BOETTCHER S BOHLE T M BOETTCHER S

RA
RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADALE Y.,
RA
FRESECCANE E., FUSIC F., FURNELLE B., KAPOFORI G., KEI M., REINOLDS
RA

RA SEKIGUCHI J., SEKOWSKA A., SEROR S.J., SEROR P., SHIN B.S., SOLDO

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE 65 KDA HYDROPHOBIC PROTEIN.
 OS HOMO SAPIENS (HUMAN).
 CC EUKARYOTA: METAEOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: PRIMATES:
 OC CATARRHINI: HOMINIDAE: HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LYMPHOMA;
 RX MEDLINE: 95155380.
 RA WILLIAMS F.M., FLINTOFF W.F.;
 RT "Isolation of a human cDNA that complements a mutant hamster cell
 defective in methotrexate uptake."
 RL J. BIOL. CHEM. 270:2987-2992(1995).
 DR EMBL: U17566; G965000;
 SO SEQUENCE 591 AA; 64852 MW; 930D15BA CRC32;

Query Match 3.4%; Score 110; DB 4; Length 591;
 Best Local Similarity 31.8%; Pred. No. 2,856-01;
 Matches 28; Conservative 22; Mismatches 29; Indels 9; Gaps 9;

DB 192 FLTFSVVALFLKPKRSLF-NRDDRGCEFSASELRMPGCGKGLHALVACG-DS 249
 DT 01-NOV-1998 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DE 243 FLEFLGWSLRR-KQVFLFEDSEFSRTONSSKDLPEF-PGHE-LPEGLSEFCIMES 299
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE GLUCOSE TRANSPORT PROTEIN.
 GN GPI.
 OS SCHISTOSOMA MANSONI (BLOOD FLUKE).
 CC EUKARYOTA: METAEOA: PLATYHELMINTHES: TREMATODA: DIGenea: STRIGEIDIDA:
 OC SCHISTOSOMATOIDEA; SCHISTOSOMATIDAE; SCHISTOSOMA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PUERTO RICAN;
 RX MEDLINE: 94140848.
 RA SKEELY P.J., KIM J.W., CUNNINGHAM J.M., SHOEMAKER C.B.;
 RT "Cloning, characterization, and functional expression of cDNAs
 encoding glucose transporter proteins from the human parasite
 Schistosoma mansoni."
 RL J. BIOL. CHEM. 269:4247-4253(1994).
 DR EMBL: L25065; G407041;
 PFAM: PF00083; sugar trf. 1.
 SO SEQUENCE 521 AA; 56815 MW; 3C17EEEB CRC32;

Query Match 3.3%; Score 107; DB 5; Length 521;
 Best Local Similarity 31.3%; Pred. No. 6,716-01;
 Matches 15; Conservative 18; Mismatches 14; Indels 1; Gaps 1;

DB 303 DVLYOYCFALGVNIVTVVSLPIERAGRTLLIMPTVSLATSL 350
 DT 01-NOV-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DE 48 DV-LQFCDFGLSVMWVVIAMARLQPVVKOVLYLLGAMLLMALQL 94
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE INTEGRAL MEMBRANE PROTEIN.
 GN EXOP.
 OS BRADYRHIZOBIUM JAPONICUM.
 CC BACTERIA: PROTEOBACTERIA: ALPHA SUBDIVISION: BRADYRHIZOBIUM GROUP:
 OC BRADYRHIZOBIUM.

FT [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JSDA 110SPC4;
 RA BECKER R.C., KOSCH K., PARINISKE M., MULLER P.;
 RT SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF039306; G2921558;
 SO SEQUENCE 786 AA; 84634 MW; 44959D2F CRC32;

Query Match 3.3%; Score 106; DB 2; Length 786;
 Best Local Similarity 40.0%; Pred. No. 8,906-01;
 Matches 24; Conservative 12; Mismatches 17; Indels 7; Gaps 6;

DB 541 YARLOE-DVHITGAILATGGGVDTLRG-WETLRPG-FPLTSLNARMDRAVTR 596
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DE 69 MARLQPVVKOVLYLLGAMLLSMA-LQDRIGLWMLGPSLFGILATATM-TVRR 125
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 92.5 KD PROTEIN C25H2.03 IN CHROMOSOME II.
 GN SPBC25H2.03.
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 CC EUKARYOTA: FUNGI: ASCOMYCOTA: ARCHIAZOSACCHAROMYCETES;
 OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
 CC SCHIZOSACCHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA DURSO G., LVE G., BOWMAN S., CHURCH C., WOOD V., BARRELL B.G.;
 RA RAJANDREAN M.A., CONNOR R.E.;
 RT SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: TO YEAST YLR386W.
 DR EMBL: Z95397; E316120;
 KW HYPOTHETICAL PROTEIN: TRANSMEMBRANE.
 FT TRANSMEM 58 78 " POTENTIAL.
 FT TRANSMEM 234 254 " POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 602 622 POTENTIAL.
 FT TRANSMEM 630 650 POTENTIAL.
 SO SEQUENCE 811 AA; 92452 MW; E1E2EB77 CRC32;

Query Match 3.3%; Score 105; DB 3; Length 811;
 Best Local Similarity 23.7%; Pred. No. 1,186+00;
 Matches 31; Conservative 39; Mismatches 52; Indels 9; Gaps 9;

DB 579 LFTPELYDLRR-LKQSAF-KLONITTYLYTAMC-HNSIAVTS-L-CLLSQNEHANL 633
 DT 01-NOV-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DE 1 ME:PPVLAIRSRVLAAYTFETMFSTFYHA-CDQPGIVFCIMDYDVLQFCDFGL 59
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DE 634 LSVAFALFENIDMLQDKLVOLIESPVFTYMRLOLEPEKPYLHKAIGLMLPOSS 693
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DE 60 MSVAVVY-IMARLQPVVKOVLYLLGAMLLSMA-LQDRIGLWMLGPSLFGILATATM 117
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DE 694 AERTLRDLOC 704
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DE 118 TVRQVRRR-HC 127

Search completed: Thu Jan 27 17:39:53 2000
 Job time: 51 secs.

Page Blank (uspto)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

BEST AVAILABLE COPY